Lecture 4. The Dual Cone and Dual Problem
For a convex cone $\mathcal{K}$, its dual cone is defined as

$$\mathcal{K}^* = \{y \mid \langle x, y \rangle \geq 0, \forall x \in \mathcal{K}\}.$$  

The inner-product can be replaced by $\langle x^T y \rangle$ if the coordinates of the vectors are given. Clearly, $\mathcal{K}^*$ is always a *closed* set, regardless if $\mathcal{K}$ is closed or not. In fact, $\mathcal{K}^*$ is also always *convex* regardless if $\mathcal{K}$ is convex or not.
In light of the dual cone, the problem of finding a lower bound on the optimal value for the standard conic optimization problem can be done as follows. Introduce the Lagrangian multiplier $y$ for the equality constraint $b - Ax = 0$, and $s$ for the conic constraint $x \in \mathcal{K}$. The objective becomes

$$c^T x + y^T (b - Ax) - s^T x = y^T b + (c - A^T y - s)^T x.$$ 

This will be a lower bound on the optimal value whenever $c - A^T y - s = 0$ and $s \in \mathcal{K}^*$. The quest to find the best lower bound leads to the dual problem for the standard conic optimization problem:

$$\text{maximize} \quad b^T y$$
$$\text{subject to} \quad A^T y + s = c$$
$$s \in \mathcal{K}^*. \quad (1)$$
As an example, for the standard SDP problem (primal),

\[
\begin{align*}
\text{minimize} & \quad C \cdot X \\
\text{subject to} & \quad A_i \cdot X = b_i, \ i = 1, \ldots, m \\
& \quad X \succeq 0,
\end{align*}
\]

its dual problem is in the form of

\[
\begin{align*}
\text{maximize} & \quad \sum_{i=1}^m b_i y_i \\
\text{subject to} & \quad \sum_{i=1}^m y_i A_i + S = C \\
& \quad S \succeq 0.
\end{align*}
\]
The so-called weak duality relationship is immediate, by the construction of the dual problem. This is shown in the next theorem. 

**Theorem 1** Let \( x \) be feasible for the primal problem, and \( (y, s) \) be feasible to its dual. Then,

\[
b^T y \leq c^T x.
\]

Moreover, if \( b^T y = c^T x \) then \( x^T s = 0 \), and they are both optimal to their own respective problems.

If a pair of primal-dual feasible solutions \( x \) and \( (y, s) \) both exist and they satisfy \( x^T s = 0 \) then we call them a pair of *complementary optimal solutions*. 

Shuzhong Zhang
A very useful theorem is the following.

**Theorem 2** Let $x$ be feasible for the primal problem, and $(y, s)$ be feasible to its dual, and $x \in \text{int} \ K$ and $s \in \text{int} \ K^\ast$. Then, the sets of optimal solutions for both problems are nonempty and compact, and any pair of primal-dual optimal solutions will be complementary to each other.
Let us spend some time now to study the structures of the dual cones. Let us start with the topological structures. If $\text{rel}(\mathcal{K})$ is denoted as the relative interior of $\mathcal{K}$, then $(\text{rel}(\mathcal{K}))^* = \mathcal{K}^*$. Let the Minkowski summation of two sets $A$ and $B$ be

\[ A + B = \{ z \mid z = x + y \text{ with } x \in A, y \in B \}. \]

It is clear that the sum of two convex sets remains convex. However, the sum of two closed convex sets may not be closed any more.
Example 1 \( \text{The set } \mathbb{R}_+ \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} + \text{SOC}(3) \text{ is convex but not closed.} \)

For instance, the point \((0, 1, 0)^T\) is not in the set, but it is in the closure of the set, since for any \(\epsilon > 0\), we may write

\[
\begin{pmatrix} \epsilon \\ 1 \\ 0 \end{pmatrix} = \frac{1}{\epsilon} \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} \frac{1}{\epsilon} + \epsilon \\ 1 \\ -\frac{1}{\epsilon} \end{pmatrix},
\]

where the second term is in \(\text{SOC}(3)\).
However, if $A$ and $B$ are both polyhedra, then $A + B$ remains a polyhedron, hence closed. It is interesting to note that the set-summation and the set-intersection operations are a duality pair:

**Proposition 1** Suppose that $\mathcal{K}_1$ and $\mathcal{K}_2$ are convex cones. It holds that

\[(\mathcal{K}_1 + \mathcal{K}_2)^* = \mathcal{K}_1^* \cap \mathcal{K}_2^* .\]

**Proof.** For $y \in (\mathcal{K}_1 + \mathcal{K}_2)^*$, it follows that $0 \leq (x_1 + x_2)^T y$ for all $x_1 \in \mathcal{K}_1$ and $x_2 \in \mathcal{K}_2$. Since $0 \in \mathcal{K}_2$, it follows that $0 \leq (x_1)^T y$ for all $x_1 \in \mathcal{K}_1$, thus $y \in \mathcal{K}_1^*$, and similarly $y \in \mathcal{K}_2^*$. Hence, $y \in \mathcal{K}_1^* \cap \mathcal{K}_2^*$.

Now, take any $y \in \mathcal{K}_1^* \cap \mathcal{K}_2^*$, and any $x_1 \in \mathcal{K}_1$ and $x_2 \in \mathcal{K}_2$. We have $(x_1 + x_2)^T y = (x_1)^T y + (x_2)^T y \geq 0$, and so $y \in (\mathcal{K}_1 + \mathcal{K}_2)^*$. This proves $(\mathcal{K}_1 + \mathcal{K}_2)^* = \mathcal{K}_1^* \cap \mathcal{K}_2^*$ as desired. \qed
The following bipolar theorem can be considered as a cornerstone for convex analysis.

**Theorem 3** *Suppose that $\mathcal{K}$ is a convex cone. Then $(\mathcal{K}^*)^* = \text{cl} \mathcal{K}$.*

**Proof.** For any $x \in \mathcal{K}$, by the definition of the dual cone, $x^T y \geq 0$ for all $y \in \mathcal{K}^*$, and so $\mathcal{K} \subseteq (\mathcal{K}^*)^*$. Therefore $\text{cl} \mathcal{K} \subseteq (\mathcal{K}^*)^*$.

The other containing relationship $(\mathcal{K}^*)^* \subseteq \text{cl} \mathcal{K}$ can be shown by contradiction. Suppose such is not true then there exists $z \in (\mathcal{K}^*)^* \setminus \text{cl} \mathcal{K}$. By the separation theorem for convex cones, there is vector $c$ such that

$$c^T z < 0, \text{ and } c^T x \geq 0 \forall x \in \text{cl} \mathcal{K}.$$ 

The last relation suggests that $c \in \mathcal{K}^*$, which contradicts with the first relation because, as $z \in (\mathcal{K}^*)^*$, it should lead to $c^T z \geq 0$, but not the opposite. This contradiction completes the proof. $\Box$
This proof suggests that the bipolar theorem for convex cone is equivalent to the separation theorem itself. Suppose that \( \mathcal{K} \) is a closed convex cone, and \( v \not\in \mathcal{K} \). Then, the bipolar theorem asserts that \( v \not\in (\mathcal{K}^*)^* \), and this implies that there is \( y \in \mathcal{K}^* \) such that \( v^T y < 0 \), and \( x^T y \geq 0 \) for any \( x \in \mathcal{K} \) because \( y \in \mathcal{K}^* \), hence a separation exists.
In case $\mathcal{K}$ is closed, the bipolar theorem is what in folklore understood as ‘the dual of the dual is the primal itself again’. Combining Theorem 3 and Proposition 1, it follows that

$$\text{cl} (\mathcal{K}_1 + \mathcal{K}_2) = (\mathcal{K}_1 \cap \mathcal{K}_2)^*.$$  

It is also useful to compute the effect of linear transformation on a convex cone in the context of dual operation. Let $A$ be an invertible matrix. Then, we have

$$(A^T \mathcal{K})^* = A^{-1} \mathcal{K}^*$$

by observing that $y \in (A^T \mathcal{K})^* \iff (A^T x)^T y \geq 0 \forall x \in \mathcal{K} \iff Ay \in \mathcal{K}^* \iff y \in A^{-1} \mathcal{K}^*.$
A convex cone is called \textit{solid} if it is full dimensional. Mathematically, \( \mathcal{K} \subseteq \mathbb{R}^n \) is solid \( \iff \) 

\[
\text{span}\, \mathcal{K} = \mathbb{R}^n, \text{ or } \mathcal{K} + (-\mathcal{K}) = \mathbb{R}^n.
\]

A convex cone is called \textit{pointed} if it does not contain any nontrivial subspace; i.e. \( \mathcal{K} \cap (-\mathcal{K}) = \{0\} \). The following result proves to be useful:

**Proposition 2** \( \text{If } \mathcal{K} \text{ is solid then } \mathcal{K}^* \text{ is pointed, and if } \mathcal{K} \text{ is pointed then } \mathcal{K}^* \text{ is solid.} \)
Proof. Suppose that $\mathcal{K}$ is solid, i.e. there is a ball $B(x_0; \delta) \subseteq \mathcal{K}$ with radius $\delta > 0$ and centered at $x_0$, and at the same time suppose that there is a line $\mathbb{R}c \in \mathcal{K}^*$, then

$$0 \leq s(x_0 + t)^T c, \forall \|t\| \leq \delta \text{ and } s \in \mathbb{R},$$

which is clearly impossible. This contradiction shows that if $\mathcal{K}$ is solid then $\mathcal{K}^*$ must be pointed. By the bipolar theorem, the second part follows by symmetry. \qed
It is easy to compute the following dual cones

- \((R_+^n)^* = R_+^n\).
- \(\text{SOC}(n)^* = \text{SOC}(n)\).
- \((S_+^{n\times n})^* = S_+^{n\times n}\), and \((H_+^{n\times n})^* = H_+^{n\times n}\).

The above standard cones are so regular that their corresponding dual cones coincide with the original cones themselves (the primal cones). In general, they are not the same. As an example, consider the following \(L_p\) cone, where \(p \geq 1\),

\[
L_p(n + 1) = \left\{ \begin{pmatrix} t \\ x \end{pmatrix} \left| \begin{array}{c} t \in \mathbb{R}, x \in \mathbb{R}^n, t \geq \|x\|_p \end{array} \right. \right\}.
\]
One can show that \((L_p(n+1))^* = L_q(n+1)\) where \(1/p + 1/q = 1\).

Now, the self-duality of \(\text{SOC}(n+1)\) is due to the fact that \(\text{SOC}(n+1) = L_2(n+1)\).

In convex analysis, the dual object for a convex function \(f(x)\) is known as the *conjugate* of \(f(x)\), defined as

\[
f^*(s) = \sup\{(-s)^T x - f(x) \mid x \in \text{dom } f\},
\]

where ‘\(\text{dom } f\)’ stands for the domain of the function \(f\). The conjugate function is also known as the *Legendre-Fenchel transformation*. 
Examples of some popular conjugate functions are:

1. For \( f(x) = \frac{1}{2}x^T Q x + b^T x \) where \( Q \succ 0 \), the conjugate function is
   \[ f^*(s) = \frac{1}{2} (s + b)^T Q^{-1} (s + b). \]

2. For \( f(x) = \sum_{i=1}^{n} e^{x_i} \), the conjugate function is
   \[ f^*(s) = -\sum_{i=1}^{n} s_i \log(-s_i) + \sum_{i=1}^{n} s_i. \]

3. For \( f(x) = \sum_{i=1}^{n} x_i \log x_i \), the conjugate function is
   \[ f^*(s) = \sum_{i=1}^{n} e^{1+s_i}. \]

4. For \( f(x) = c^T x \), the conjugate function is \( f^*(s) = 0 \) for \( s = -c \)
   and \( f^*(s) = +\infty \) when \( s \neq -c \).

5. For \( f(x) = -\sum_{i=1}^{n} \log x_i \), the conjugate function is
   \[ f^*(s) = -n - \sum_{i=1}^{n} \log s_i. \]

6. For \( f(x) = \max_{1 \leq i \leq n} x_i \), the conjugate function is \( f^*(s) = 0 \) if
   \( \sum_{i=1}^{n} s_i \leq -1 \), and \( f^*(s) = +\infty \) otherwise.
If \( f(x) \) is strictly convex and differentiable, then \( f^*(s) \) is also strictly convex and differentiable. The famous bi-conjugate theorem states that \( f^{**} = \text{cl} \ f \). There is certainly a relationship between the two dual objects: the dual cone and the conjugate function. Let us now set out to find the exact relationship. First of all, suppose that \( f \) is a convex function, mapping from its domain in \( \mathbb{R}^n \) to \( \mathbb{R} \). Its epigraphy is a set in \( \mathbb{R}^{n+1} \) defined as

\[
\text{epi} \ (f) := \left\{ \begin{pmatrix} t \\ x \end{pmatrix} \mid t \geq f(x), \ t \in \mathbb{R}, \ x \in \text{dom} \ (f) \right\}.
\]
It is well known that $f$ is a convex function if and only if epi ($f$) is a convex set. It turns out that $H(\text{epi}(f^*))$ is the dual of $H(\text{epi}(f))$, after taking the closure and reorienting itself. To be precise, we have:

**Theorem 4** Suppose that $f$ is a convex function, and

\[
\mathcal{K} := \text{cl} \left\{ \begin{pmatrix} p \\ q \\ x \end{pmatrix} \mid p > 0, q - pf(x/p) \geq 0 \right\}.
\]

Then,

\[
\mathcal{K}^* = \text{cl} \left\{ \begin{pmatrix} u \\ v \\ s \end{pmatrix} \mid v > 0, u - vf^*(s/v) \geq 0 \right\}.
\]
Let $L_2 := \begin{bmatrix} 0, & 1 \\ 1, & 0 \end{bmatrix}$ and $L := L_2 \oplus I_n = [L_2, 0_{2,n}; 0_{n,2}, I_n]$. The effect of applying $L$ to $x$ is to swap the position of $x_1$ and $x_2$. Equivalently, Theorem 4 can be stated as

$$(H(\text{epi}(f)))^* = \text{cl}(L(H(\text{epi}(f^*))))$$.
Another related concept is the so-called polar set. Suppose $S \subset \mathbb{R}^n$. Then its polar set is defined as

$$S^\circ = \{ y \mid (-y)^T x \leq 1, \forall x \in S \}.$$  

Clearly, $S^\circ$ is always closed and convex, and it always contains the origin.

If $S$ is a cone, then $S^\circ = S^*$. 

One can verify that

$$S^\circ = \left\{ y \mid \begin{pmatrix} 1 \\ y \end{pmatrix} \in H(S)^* \right\}.$$
Some basic properties regarding the polar set immediately follow:

- If $A \subseteq B$ then $A^\circ \supseteq B^\circ$;
- $(\cup_{i \in I} S_i)^\circ = \cap_{i \in I} S_i^\circ$;
- If $S = \text{conv}(v_1, \ldots, v_m)$ then
  \[
  S^\circ = \{ x \mid -v_i^T x \leq 1, \ i = 1, 2, \ldots, m \}.
  \]
  In other words, the polar of a polytope is a polyhedron.
- It always holds $S \subseteq S^{\circ \circ}$.
By Theorem 3, we also have

**Theorem 5** *If $S$ is a convex set, then*

$$S^{\circ\circ} = \text{cl } S.$$  

The above is evidently equivalent to Theorem 3. They can both be referred to as the bipolar theorem.
Let us return to primal-dual conic optimization models.

In linear programming, it is known that whenever the primal and the dual problems are both feasible then complementary solutions must exist. This important property, however, is lost, for a general conic optimization model, as the following example shows.

Consider

$$\begin{align*}
\text{minimize} & \quad x_{11} \\
\text{subject to} & \quad x_{12} + x_{21} = 1, \\
 & \quad \begin{bmatrix} x_{11}, & x_{12} \\
 x_{21}, & x_{22} \end{bmatrix} \succeq 0.
\end{align*}$$

This problem is clearly feasible. However, there is no attainable optimal solution.
Its dual problem is

\[
\begin{align*}
\text{maximize} & \quad y \\
\text{subject to} & \quad y \begin{bmatrix} 0, & 1 \\ 1, & 0 \end{bmatrix} + \begin{bmatrix} s_{11}, & s_{12} \\ s_{21}, & s_{22} \end{bmatrix} = \begin{bmatrix} 1, & 0 \\ 0, & 0 \end{bmatrix}, \\
& \begin{bmatrix} s_{11}, & s_{12} \\ s_{21}, & s_{22} \end{bmatrix} \succeq 0,
\end{align*}
\]

or, more explicitly,

\[
\begin{align*}
\text{maximize} & \quad y \\
\text{subject to} & \quad \begin{bmatrix} 1, & -y \\ -y, & 0 \end{bmatrix} \succeq 0.
\end{align*}
\]

The dual problem is feasible and has an optimal solution; however, a pair of primal-dual complementary optimal solutions fails to exist, although both primal and dual problems are feasible.
One may guess that the unfortunate situation was caused by the non-attainable optimal solution. Next example shows that even if both primal and dual problems are nicely feasible with attainable optimal solutions, there might still be a positive duality gap, meaning that the complementarity condition does not hold.

\[
\begin{align*}
\text{minimize} & \quad -x_{12} - x_{21} + 2x_{33} \\
\text{subject to} & \quad x_{12} + x_{21} + x_{33} = 1, \\
& \quad x_{22} = 0, \\
& \quad \begin{bmatrix}
  x_{11}, & x_{12}, & x_{13} \\
  x_{21}, & x_{22}, & x_{23} \\
  x_{31}, & x_{32}, & x_{33}
\end{bmatrix} \succeq 0.
\end{align*}
\]
The dual of this problem is

\[
\text{maximize } y_1
\]

\[
\text{subject to } \begin{bmatrix} 0, & -1 - y_1, & 0 \\ -1 - y_1, & -y_2, & 0 \\ 0, & 0, & 2 - y_1 \end{bmatrix} \succeq 0.
\]

The primal problem has an optimal solution, e.g.

\[
X^* = \begin{bmatrix} 1, & 0, & 0 \\ 0, & 0, & 0 \\ 0, & 0, & 1 \end{bmatrix},
\]

with the optimal value equal to 2. The dual problem also is feasible, and has an optimal solution e.g.

\((y_1^*, y_2^*) = (-1, -1),\) with optimal value equal to \(-1.\) Both problems are feasible and have attainable optimal solutions; however, their optimal solutions are not complementary to each other, and there is a positive duality gap: \(c^T x^* - b^T y^* = (x^*)^T s^* > 0.\)
Key References:

